Wanous, M. (2017). A Neural Network Mark-up Estimation Model for Syrian Contractors. *International Journal of Architecture, Engineering and Construction*, *6*(1), 41-49. [5]. https://doi.org/10.7492/IJAEC.2017.005

Peer reviewed version

Link to published version (if available):
10.7492/IJAEC.2017.005

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via International Association for Sustainable Development and Management (IASDM) and North China University of Water Resources and Electric Power at DOI: 10.7492/IJAEC.2017.005. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

A Neural Network Mark-up Estimation Model for Syrian Contractors

WANOUS, Mohammed* and BOUSSABAINE, A. Halim**

** Senior Lecturer in Engineering Management, Department of Civil Engineering, University of Bristol, Queen's Building, University Walk, Bristol, BS8 1TR, UK. Email:* m.wanous@bristol.ac.uk

*** Professor of Project Management, Faculty of Engineering & IT, The British University in Dubai, Dubai, United Arab Emirates.*

Abstract

One of the most important decisions that have to be made by construction contractors is how much to mark-up the estimated cost of a new project. The main objectives of this paper are to model the relationship between mark-up estimation and the key factors affecting it and to compare the application of regression analysis and neural network techniques on the mark-up decision making process in order to find which technique is more reliable in terms of accuracy and robustness. The most influential mark-up factors were identified through a formal questionnaire survey conducted among Syrian contractors. Subsequently, data on one hundred and eleven real-life bidding situations was collected from Syria. Ninety-six of these projects were used to develop linear, non-linear regression and neural network mark-up models. The remaining fifteen projects were randomly held-back for validating the developed models. The neural network model proved to be robust and more accurate than the regression models. Although this study was carried out in the context of the Syrian construction industry, the methodology and the findings have much broader geographical applicability.

Keywords: Mark-up size Criteria, Regression Analysis, Neural Networks, Modelling, Syria.

Introduction

Pricing a bid for a new project normally consists of a two-stage process comprising baseline estimate and mark-up (Drew and Skitmore, 1997). Generally, the components of mark-up include profit, risk contingencies, and recovery of general overheads. The actual process of how contractors determine their bidding price, which includes mark up, is not clearly articulated in the literature (Laryea and Hughes, 2011). Selection of an appropriate mark-up for a new construction project is a very complex decision-making process (Ahmad, 1988). Different bidders apply different mark-up policies (Drew and Skitmore, 1997). In practice, most Syrian contractors adjust productivity factors or add contingencies based on the risk of each item being estimated, include the general overheads in the indirect cost and then apply a standard mark-up to the total cost estimate to cover profit and any unforeseen risks not allowed for in the cost estimate. Therefore, the development of an effective decision-support model for setting a suitable mark-up size for new construction projects can yield significant benefits, especially for contractors who do not have much experience in this . However, the aim of modelling the mark-up decision making process, is not to replace decision makers but for such models to be used in training exercises and to provide broad guidelines for senior management. Also, mark-up models help contractors to attain a reasonable degree of consistency and to check for mistakes. During the last fifty years, numerous models have been developed for the mark-up selection process. Most of these models remain in the academic domain and have not found their way into the practical world for many reasons, such as (Drew and Skitmore, 1997; Ahmad, 1988; Wanous et al. 2000):

1- Over-simplicity of assumptions in many models making them unrepresentative of the real world;
2- Most contractors being unwilling to struggle with sophisticated mathematical models; and
3- Most of these bidding models not taking into account that contractors might have other objectives rather than maximising their expected profit. These might include minimising

expected loss, minimising profit of competitors, gaining a strategic market position, or maintaining a certain level of workload.

The mark-up decision is often based on heuristic techniques, i.e. experience, subjective judgement and intuition of the decision maker (Wanous et al. 1998; Couzens et al. 1996). This paper studies the ability of regression analysis and neural network techniques to capture the basic intuitive heuristic techniques used in real life by contractors when making mark-up decisions. Through a questionnaire survey the main factors that influence this decision were identified and ranked according to their importance to contractors operating in Syria. Only the most influential factors were then considered in developing a template form for collecting data on real life bidding situations. The assessment of the mark-up factors and the accompanying mark-up values were collected for one hundred and eleven bidding case studies. These bidding situations consisted of building projects (48.65%), roads (18.92%), pipelines (29.73%) and dams (2.7%). Fifteen projects were randomly selected from the original sample and held back for testing. The remaining ninety-six projects were used to develop regression and neural network models.

Previous Studies

The literature contains a great number of theoretical bidding models based on the expected monetary value and the expected utility value. The first probability-based bidding model was proposed by Freidman (Friedman, 1956). Many researchers have discussed the validity and practicality of such probabilistic monetary value models (Gates, 1976; Benjamin and Meador, 1979; Ioannou, 1988). The most important points of their debates are the over-simplicity of the models' assumptions and the necessity of historical data about past projects and competitors. It is unlikely that any given contractor could acquire enough data to be able to develop the probability distribution of known competitors' bid-to-cost ratios that are needed for these models. Nevertheless, these models have made a significant contribution in formalising the mark-up decision-making process.

Ahmad and Minkarha (1987) developed a multi-dimensional utility model. They defined three utility functions for the contractor's preference structure, attitude towards loss, and the general overheads. The main advantage of this model is its ability to consider a contractor's preference structure and to handle multi-criteria decision-making problems. Also, it enables subjective judgements to be used to assess the relative importance of the considered criteria. However, the necessity of historical data, which is usually difficult to obtain undermines the applicability of this model to actual bidding situations.

Dikmen et al. (2007) collected from Turkish contractors 41 factors influencing bid mark-up under three categories; risk, opportunity and competition. They developed a linear mark-up model using utility functions for risk, opportunities and competition with an attempt to take different strategies and preferences of contractors into account. Broemser (1968) proposed two bidding models (single bid model and sequential bid model) that consider the effect of other factors apart from maximising the expected profit. These factors include project size, risk, proportion of the job to be subcontracted, and the number of competitors. A linear regression performed on data collected from contractors in California on past bidding situations and their outcomes predicted the effect of each of these factors on the mark-up. The results of the regression analysis revealed that the probability of winning is not a function of the number of competitors as assumed by the previous models. Very few qualitative approaches, which study how bidding decisions are made in practice, have been carried out. Gates (1983) suggested a non-mathematical bidding strategy based on the Delphi technique, designated as the ESPE (Expert Subjective Pragmatic Estimate). In this model, the range of competitors' possible low bids is estimated and another estimate is made for the company's range and distribution of the possible low bids. The two sets are then compared to select the most appropriate bid. This is done by a group of experts who, through an iterative process, estimate the best bid. Ahmad and Minkarah (1988) conducted a questionnaire survey to uncover the factors that characterise the bidding decision-making process in the United States. Degree of hazard, degree of difficulty, and uncertainty in cost estimate were the top three mark-up factors. Shash (1993) identified, through a modified version of the same questionnaire used by Ahmad and Minkarah (1988), fifty five factors that characterise the mark-up size decision in the UK. The need for work, number of competitors tendering and experience on similar projects were identified as the top three factors that affect the mark-up decision. Egemen and Mohamed (2008) proposed a knowledge-based system called SCBMD to deal with different bidding situations and help contractors in making bid/no bid and mark-up selection decisions. ElSawy, et al (2011) proposed a parametric Artificial Neural Network cost-estimating model for site overhead in Egypt based on 52 real-life construction projects. Polat, et al (2015) developed a mark-up size estimation model for international construction projects using the integration of AHP and Regression Analysis techniques. Gaarslev (1991) used the data that was used by Broemser (1968) to develop a neural network mark-up model. Surprisingly, it was concluded that the neural network model does not produce valuable predictions and Broemser's regression model produces better results with minor effort. This result seems unconvincing because neural network technology itself is somehow an automatic regression technique. Additionally, neural networks allow a higher degree of freedom to accommodate any non-linearity in the model being developed. The current work will re-examine this situation by comparing the application of neural network and regression analysis on the mark-up selection process.

Key Mark-up Factors

A formal questionnaire, as shown in Appendix 1, was prepared to seek the opinions of Syrian general contractors about the importance of factors that affect their "bid/ no bid" and mark-up decisions. The questionnaire started with general questions about the contractor such as the typical project type(s)/size the contractor usually deals with and method used in making "bid/no bid" and mark-up decisions. Then, it listed thirty-eight factors that were assumed to influence bidding decisions in Syria. Contractors were prompted to add any missing factors and to express their opinions about the importance of each factor by circling the appropriate score from 0 to 6 (where: 0 means no importance at all; and 6 means extreme importance). Analysis the questionnaire responses revealed the relative importance of thirty five bidding factors as considered by Syrian contractors. These factors were ranked according to their influence on the mark-up size using an index called the importance index (I). The following equation was used to produce this index:

$$I_j = M_j / 6 \qquad \text{....(1)}$$

Where:

6 is the maximum possible score as set in the questionnaire.

I_j is the importance index (0 to 1) of factor j in selecting the mark-up size; and

M_j is the mean importance score of factor j. M_j is produced using the following formula (Medhi, 1992):

$$M_j = \frac{\sum_{i=0}^{i=6} (s_{ij} * n_{ij})}{N_j} \qquad \text{....(2)}$$

Where:

s_{ij} : score between 0 and 6 given to factor j by each contractor;

n_{ij}: number of contractors who scored factor j by s_{ij};

N_j: number of contractors who gave a score to factor j .

Only nineteen factors, which have an importance index (I) equal to or greater than 50%, were considered (see Table 1) to develop a simple template form to collect data on real life bidding situations. In this standard template, respondents were asked to provide the actual mark-up (as a percentage of the total estimated cost) and their subjective assessments in current or recent bidding situations in terms of the nineteen factors listed in the form. One hundred and eleven forms were filled in and returned. Fifteen cases were randomly selected and reserved for the validation process. The remaining ninety-six bidding situations were used to study the correlation between the contractors' assessments of the mark-up factors and the actual mark-up values. Table 1 shows these factors ranked according to their relationship with the actual mark-up size expressed by the absolute correlation coefficient |r|.

Table 1: Selection of the most influential factors

No.	**The most influential mark-up factors**	**I**	**r**	**T**	**Signif T**
1	Risks expected	0.887	+0.711	4.737	0.0000*
2	Availability of equipment owned by the contractor	0.653	-0.636	-1.728	0.0874*
3	Confidence in the cost estimate	0.752	-0.631	-2.080	0.0404*
4	Availability of materials required	0.783	-0.619	-2.595	0.0111*
5	Competence of the expected competitors	0.837	-0.614	-3.366	0.0011*
6	Degree of buildability	0.772	-0.596	-1.367	0.1751
7	Expected degree of competition (number of competitors)	0.825	-0.577	-0.876	0.3834
8	Way of construction (mechanically/ manually)	0.625	-0.544	-2.114	0.0374*
9	Rigidity of specifications	0.788	+0.533	0.379	0.7057
10	Site clearance of obstructions	0.588	-0.528	-0.692	0.4906
11	Site accessibility	0.615	-0.514	-0.293	0.7701
12	Public objection	0.645	+0.208	-	-
13	Remoteness of the project location	0.603	+0.199	-	-
14	Experience on similar projects	0.517	-0.147	-	-
15	Availability of skilled labour	0.560	-0.088	-	-
16	Project size	0.655	-0.021	-	-
17	Current workload	0.547	+0.010	-	-
18	Sufficiency of the project duration	0.575	-0.007	-	-
19	Availability of equipment required	0.658	+0.005	-	-

I is the importance index
r is the correlation coefficient between the mark-up and the mark-up factors
* Denoting the considered factors in the linear regression model

The factors that have marginal correlation with the mark-up ($|r| < 0.5$) were omitted. The remaining eleven factors were considered in developing the regression and the artificial neural network (ANN) mark-up models as explained in the following sections.

Mark-up Selection: A Linear Regression Analysis Approach

The eleven factors selected in the previous section were used to develop a linear regression equation that best fits the modelling sample (ninety-six bidding situations). The SPSS statistical package was used to perform various methods of linear regression (Enter, Stepwise, Forward, and Backward). The Forward and the Backward regression methods produced the same model, which considers only six input variables. The adjusted R squared of this model (0.713) is higher than that of the other linear models. Therefore, it was selected as the best-fit linear regression model. The fifth column of Table 1 shows the T values of the mark-up factors produced by the Forward regression method.

Asterisks in the last column of this table denote the factors that are considered. The selected linear model is given in the following equation:

$$\text{Mark-up} = 0.221841 - 0.006842*F_1 - 0.005385*F_2 - 0.00314*F_3 + 0.00677*F_4 - 0.002333*F_5 - 0.00816*F_8 \quad (3)$$

The linearity assumed in this model might or might not be true. Thus, a non-linear regression approach was implemented to develop the best possible non-linear mark-up model as explained in the next section.

Mark-up Selection: A Non-Linear Regression Approach

The development of a non-linear regression model is basically an iterative trial and error process. As no standard procedures are available for developing non-linear regression models, an attempt was made in this study to systemise this process as summarised below:

1. The actual mark-up values in the modelling sample were plotted against the contractors' assessments of each individual factor using scatter diagrams.
2. The best trend line along with its equation and R squared value were produced for each factor. For example, Fig. 1 shows the relation between mark-up and the first factor (risks expected).

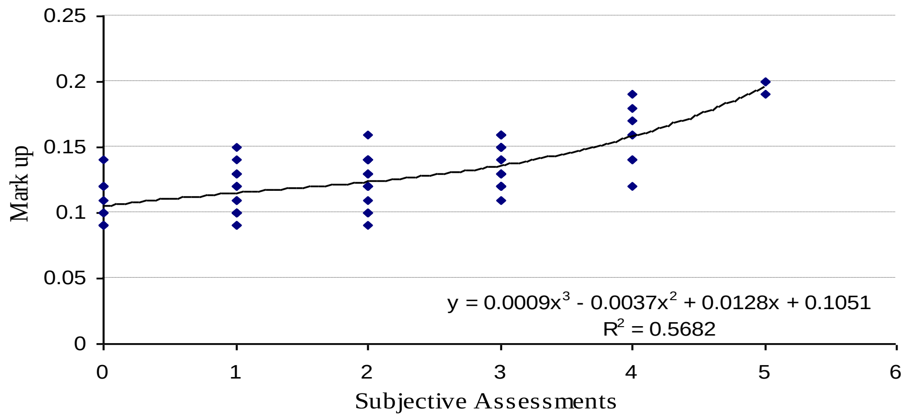

Fig.1: The best-fit correlation between the mark up and the "Risks expected"

The equations and the R squared values produced for the eleven individual factors are shown in Table 2.

Table 2: The best-fit non-linear models considering individual mark-up factors

Factor	Non-linear equation	R-Squared
F_1	$M= 0.1051+0.0128 *F_1-0.0037* F_1^2+0.009* F_1^3$	0.5682
F_2	$M= 0.1739-0.0247*F_2+0.0046* F_2^2-0.0005* F_2^3$	0.4321
F_3	$M= 0.2494*EXP(-0.1652* F_3)$	0.4281
F_4	$M= 0.5768-0.2544*F_4+0.0496* F_4^2-0.0034* F_4^3$	0.4240
F_5	$M= 0.8985-0.4789*F_5+0.1009* F_5^2-0.0072* F_5^3$	0.4200
F_6	$M= 0.2512-0.0747*F_6+0.0168* F_6^2-0.0015* F_6^3-0.000021*F_6^4$	0.4050
F_7	$M= 0.4906-0.2132*F_7+0.044* F_7^2-0.0033* F_7^3+0.000014* F_7^4-0.000012*F_7^5$	0.3304
F_8	$M= 0.1905+0.0222 *F_8-0.029* F_8^2+0.0067* F_8^3-0.0005* F_8^4-0.000021* F_8^5$	0.3206
F_9	$M=-0.0214+0.0639*F_9+0.0063* F_9^2-0.0061* F_9^3+0.0007*F_9^4$	0.3204
F_{10}	$M= 0.182*EXP (-0.0962* F_{10})$	0.3060
F_{11}	$M= 0.5233-0.4101*F_{11}+0.1703* F_{11}^2-0.0318* F_{11}^3+0.0022* F_{11}^4-0.000041*F_{11}^5$	0.2950

3. Individual equations provided a range of non-linear parameters to choose from during the development of non-linear models. Starting with the equation using the first factor (because it has the highest R square), parameters from the second equations were added one parameter at a time. After adding a new parameter, the equation was experimented with using the SPSS package and the resulting R square was recorded. When adding a parameter reduced the R square value, it was omitted.

In this way, more than seventy equations were examined before developing the final non-linear mark-up model shown below:

$$
\begin{aligned}
\text{Mark-up} = {} & -9.441112279-0.00628225*F_1+0.003808863*F_1\text{^}2-\\
& 0.000284558*F_1\text{^}3-0.288816319*F_2+0.06137285*F_2\text{^}2-\\
& 0.004294553*F_2\text{^}3-0.002802286*\mathrm{EXP}(0.419293361*F_3)-\\
& 0.165006062*F_4+0.03508057*F_4\text{^}2-0.002483252*F_4\text{^}3-\\
& 0.011653475*F_5+0.00396265*F_5\text{^}2-0.000437456*F_5\text{^}3+\\
& 0.510946414*F_6-0.186452872*F_6\text{^}2+0.029107653*F_6\text{^}3-\\
& 0.001653161*F_6\text{^}4+12.377261191*F_7-6.887296762*F_7\text{^}2+\\
& 1.846476669*F_7\text{^}3-0.23931915*F_7\text{^}4+0.012030553*F_7\text{^}5-\\
& 0.094035992*F_8+0.109280611*F_8\text{^}2-0.05144907*F_8\text{^}3+\\
& 0.010500718*F_8\text{^}4-0.000774616*F_8\text{^}5+0.509122872*F_9-\\
& 0.199398855*F_9\text{^}2+0.032903452*F_9\text{^}3-0.001948869*F_9\text{^}4-\\
& 10.98309836*\mathrm{EXP}(0.000259526*F_{10})+17.130708662*F_{11}-\\
& 9.522532569*F_{11}\text{^}2+2.548534709*F_{11}\text{^}3-0.32951393*F_{11}\text{^}4+\\
& 0.016513874*F_{11}\text{^}5 \qquad (4)
\end{aligned}
$$

The adjusted R-squared and the sum of squared residuals of this model are 0.81 and 0.0108 respectively. This shows that the non-linear model fits the modelling data better than the linear model developed in the previous section. However, non-linear regression models are known to be unstable, i.e. small changes in the input space might cause large changes in the output space. Also, extreme inputs might produce unrealistic outputs (e.g. negative mark-up). Therefore, the stability of the non-linear model is examined in the following section.

Stability Analysis of Non-linear Regression Model

To test the stability of the non-linear model, its sensitivity to variation in the inputs was examined. The model inputs are subjective assessments. The outputs were recorded while changing the assessment of the first factor (F1) and setting the assessments of the remaining factors to medium. The same process was repeated for all the factors. Table 3 shows the outputs produced by the model for different assessments of the input factors.

The shaded cells of this table show that the model will recommend mark-up values with excessive departure from the usual practice if certain factors were assigned extreme scores.

Table 3: Sensitivity of the non-linear model to changes in its input variables

Factors	**Assessments**						
	0	**1**	**2**	**(3)**	**4**	**5**	**6**
F1	0.1761	0.1734	0.1765	0.1839	0.1938	0.2044	0.2141
F2	0.6139	0.3822	0.2475	0.1839	0.1658	0.1673	0.1628
F3	0.1910	0.1895	0.1873	0.1839	0.1788	0.1710	0.1591
F4	0.4302	0.2978	0.2207	0.1839	0.1726	0.1718	0.1667
F5	0.1950	0.1869	0.1841	0.1839	0.1838	0.1811	0.1733
F6	-0.3229	0.0291	0.1596	0.1839	0.1774	0.1758	0.1753
F7	-8.3557	-1.2465	0.1774	0.1839	0.1847	0.1788	0.1962
F8	0.1921	0.1656	0.1727	0.1839	0.1666	0.1650	0.0345
F9	-0.2794	0.0613	0.1733	0.1839	0.1736	0.1761	0.1784
F10	0.1925	0.1896	0.1868	0.1839	0.1810	0.1792	0.1753
F11	-11.638	-1.7944	0.1777	0.1839	0.1851	0.1786	0.1803

The modelling data does not contain any case where similar extreme scores were assigned to these factors. This might be the reason for the model being unable to give reasonable recommendations in such cases. Although, it can be justified in some special situations, negative or excessively high mark-up is uncommon in actual practice. Additionally, it may be observed in Table 3 that small variations in certain factors will cause big variation in the model output, which undermine the model's stability. Similar analysis was performed on the linear regression model. The linear model proved to be more stable and does not produce similar unrealistic outputs for extreme inputs.

The next step is to examine if the Artificial Neural Network technique can produce a better model. The suitability of the ANN technique to mark-up estimation has been supported by many authors (Moselhi et al. 1991; Boussabaine, 1995). Although the ANN technique does not guarantee the best model, it makes the development much easier because it can correlate outputs and inputs automatically when built with adequate data. Whereas, the non-linear regression analysis technique requires that the user provides the equations before testing them on available samples. The possibility of developing a better mark-up model and to compare the performance of the ANN with regression techniques is therefore investigated in the following section.

The Development of a Neural Network Mark-up Model

The same ninety-six projects used in developing the regression models were used to develop an ANN model with the same input factors. The development procedure adopted is explained in the following sub-sections.

1- Initial Design Assumptions

The mark-up factors were considered as the input variables of the ANN mark-up models. The simplest topology was adopted for the initial model (M1) as a starting point. The input buffer contained eleven nodes (one for each input factor) fully connected to the output layer, which contained only one processing element (PE) for the only output (mark-up percentage). The "normalised cumulative delta" learning rule and the sigmoid transfer function were used. The other parameters were set to their default values selected by the development software (NeuralWorks). The initial weights were automatically set to random small numbers between (-0.5) and (+0.5).

2- Training

The back propagation learning algorithm was used to modify initial connection weights. A fixed number of training iterations (50000) was used in this stage. When the learning counter reaches this limit, the learning was automatically ceased. The ability of model (M1) to explain the variance in the training data after 50000 iterations was presented by its training diagnostic instruments (RMS_{train}=0.0566 and R^2_{train}=0.8413). These values were recorded (Table 4).

Table 4: Selection of the best ANN model

One output (Mark-up percentage)											
M Net	No. Inputs	No. of Hidden Layers	Node s In H.L 1	Nodes In H.L 2	Iteration	L.R.	T.F.	Training		Testing	
								RMS	R^2	RMS	R^2
1	11	0	0	0	50000	N-C-D	Seigmoid	0.0566	0.8413	0.0538	0.9113
2		1	5	0				0.0592	0.9436	0.0571	0.8989
3			10					0.0659	0.6648	0.0579	0.8962
4			15					0.0630	0.7442	0.0569	0.9005
5			20					0.0586	0.7951	0.0558	0.9039
6			25					0.0802	0.8849	0.0567	0.9009
7			30					0.0752	0.7744	0.0562	0.9026
8		2	5	1				0.1117	0.4740	0.1084	0.7760
9				2				0.0690	0.9153	0.0688	0.8485
10				5				0.0531	0.5794	0.0633	0.8737
11				10				0.0746	0.6571	0.0648	0.8668
12		2	10	1				0.0879	0.7907	0.0678	0.8544
13				2				0.0648	0.8530	0.0669	0.8586
14				5				0.0755	0.8776	0.0588	0.8923
15				10				0.0800	0.8457	0.0596	0.8889
16		2	15	1				0.0903	0.7769	0.0677	0.8544
17				2				0.0612	0.8292	0.0611	0.8822
18				5				0.0780	0.8329	0.0610	0.883
19				10				0.0711	0.9092	0.0577	0.8962
20		1	5	0	50000	D-R	Seigmoid	0.0580	0.9543	0.0566	0.9166
21						ExtDBD		0.587	0.9501	0.0521	0.9163
22						QP		0.0613	0.9378	0.0589	0.8935
23						MP		0.0919	0.6370	0.1169	0.8605
24						D-B-D		0.0665	0.9298	0.0731	0.8782
25		1	5	0	50000	D-R	Linear	0.1732	0.9212	0.1463	0.9068
26							TanH	0.1375	0.9635	0.2291	0.7637
27							DNNA	0.0639	0.9275	0.0697	0.8440
28							Sine	0.1290	0.9602	0.1891	0.8505

29						E 5			0.0523	0.9929	0.0519	0.9166
30						E 10			0.0620	0.8193	0.0519	0.9166
31		1	5	0	50000	E 20			0.0611	0.9403	0.0519	0.9166
32						E 25	D-R	Seigmoid	0.0571	0.8539	0.0519	0.9166
33						E 30			0.0754	0.8518	0.0519	0.9166
34					52900				0.0339	0.9852	0.0517	0.9174
35		1	5	0	65730				0.0261	0.9572	0.0517	0.9173
36					67530				0.0303	0.9982	0.0518	0.9170

3- Testing

The fifteen bidding situations reserved for the validation process were used to examine the generalisation capability of model (M1) after training. The contractors' assessments of these situations were presented to this model. The outputs produced were compared to the actual mark-up values and the software used provided two measures of the test result. These measures (RMS_{test} = 0.0.0538 and R^2_{test} = 0.9113) were recorded in Table 4.

4- Model Selection

Starting with the initial model "M1", many combinations of different number of hidden layers with different number of processing elements (PEs) were tried (M2 to M19). Training and testing results were recorded in Table 4.

Model M2 (one hidden layer containing five PEs) produced the best result up to this stage. The structure of M2 was therefore adopted in subsequent models with different learning rules (M20 to M24). The "delta rule" learning mechanism performed better that the others.

In models M25 to M28, different transfer functions were tested. The sigmoid proved more suitable compared with other transfer functions. Thus, it was used in all subsequent models. The “delta rule” learning mechanism does not use the epoch size (E) when updating the connection weights. Nevertheless, different epoch sizes were tested (in models M29 to M33) because the development software uses the epoch size to calculate the RMS and R^2 measures during the training process. The epoch size (E = 5) helped to get the lowest RMS and the highest R^2 values. Thus it was used in all subsequent models.

Model (M29) showed the best performance. Therefore, it was selected and many attempts were made to improve it by more training (M34 to M36).

The training and testing results in terms of RMS and R^2 for all models shown in Table 4 indicate that model M36 has the best overall performance. Thus, this model was selected as the best ANN mark-up model. It is composed of the following layers as illustrated in Fig. 2:

1. Input layer (*I*) containing eleven nodes for the eleven mark-up variables and a bias node (f0), the input of which is always equal to one. The input nodes are fully connected to the next layer. The bias node is connected to all the subsequent layers;
2. Hidden layer (*J*) containing five processing elements with sigmoid transfer function. All these PEs are fully connected to the output layer; and
3. Output layer (*O*) containing one output processing element with a sigmoid transfer function.

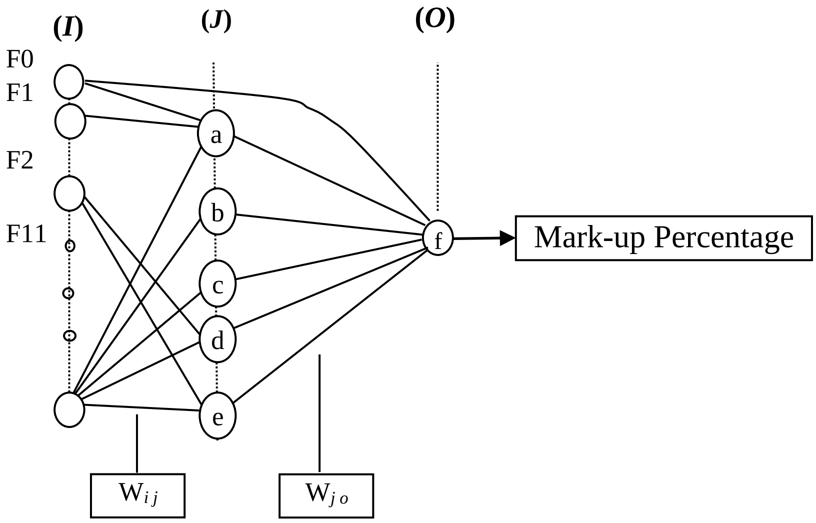

Fig. 2: Structure of the final ANN mark-up model

The following section examines the robustness of this model.

Stability Analysis of the ANN Model

The outputs produced by the ANN model were recorded in Table 5 while changing the assessment of the first factor (F1) between extremely low (0) to extremely high (6). Meanwhile, all other input factors were set to the mid-point score, i.e. medium (3). The same process was repeated for all factors.

Table 5: Sensitivity of the neural network model to changes in its input variables

Factors	Assessments						
	0	1	2	(3)	4	5	6
F1	0.1320	0.1420	0.1460	0.1541	0.1625	0.1682	0.1751
F2	0.1620	0.1632	0.1528	0.1541	0.1475	0.1408	0.1389
F3	0.1685	0.1631	0.1558	0.1541	0.1469	0.1336	0.1312
F4	0.1725	0.1709	0.1631	0.1541	0.1436	0.1375	0.1311
F5	0.1241	0.1615	0.1584	0.1541	0.1491	0.1436	0.1367
F6	0.1658	0.1575	0.1559	0.1541	0.1474	0.1429	0.1381
F7	0.1712	0.1696	0.1569	0.1541	0.1464	0.1432	0.1342
F8	0.1441	0.1455	0.1524	0.1541	0.1549	0.1579	0.1632
F9	0.1569	0.1557	0.1549	0.1541	0.1536	0.1520	0.1503
F10	0.1689	0.1621	0.1587	0.1541	0.1405	0.1309	0.1301
F11	0.1704	0.1658	0.1565	0.1541	0.1452	0.1352	0.1324

Table 5 shows that:

1. Extreme values of any input variable does not cause the ANN model to produce unrealistic mark-up recommendations; and
2. Small changes in any input variable do not cause large changes, i.e. steps, in the output of this model.

Thus, it is concluded that the ANN mark-up model provided a successful solution for the lack of stability, which in comparison undermines the reliability of the non-linear regression model. However, all the developed models need to be validated against new bidding situations before stating the superiority of any particular model. This is explained in the following section.

Final Testing and Validation

The linear, non-linear, and the ANN models were used to predict the mark-up values of the same fifteen bidding situations reserved for validation. Table 6 shows the actual mark-up values, the predicted values, Errors (E), Absolute Percentage Errors (APE) and the Mean Absolute Percentage Error (MAPE) for the three models.

Table 6: Comparison between linear, non-linear regression and ANN mark-up models

Test Projects	Actual Mark-up	Linear Model			Non-linear Model			Neural network		
		Predicted	E	APE	Predicted	E	APE	Predicted	E	APE
1	0.12	0.132	-0.012	8.84	0.118	0.002	1.92	0.123	-0.003	2.739
2	0.14	0.130	0.010	7.58	0.128	0.012	9.21	0.125	0.015	10.392
3	0.15	0.120	0.030	24.55	0.126	0.024	18.90	0.134	0.016	10.869
4	0.13	0.152	-0.022	14.56	0.138	-0.008	6.07	0.156	-0.026	20.254
5	0.18	0.160	0.020	12.34	0.170	0.010	6.22	0.168	0.012	6.701
6	0.15	0.128	0.022	17.62	0.124	0.026	21.08	0.132	0.018	11.730
7	0.18	0.172	0.008	4.60	0.181	-0.001	0.66	0.174	0.006	3.416
8	0.16	0.138	0.022	16.16	0.126	0.034	27.31	0.144	0.016	9.914
9	0.12	0.111	0.009	8.17	0.109	0.011	10.15	0.111	0.009	7.162
10	0.11	0.114	-0.004	3.60	0.116	-0.006	5.35	0.116	-0.006	5.888
11	0.10	0.096	0.004	3.81	0.103	-0.003	3.33	0.097	0.003	2.931
12	0.09	0.112	-0.022	19.36	0.110	-0.020	17.89	0.105	-0.015	16.427
13	0.13	0.129	0.001	0.70	0.117	0.013	10.92	0.120	0.010	7.400
14	0.15	0.136	0.014	10.17	0.145	0.005	3.27	0.142	0.008	5.271
15	0.11	0.108	0.002	2.27	0.113	-0.003	2.17	0.116	-0.006	5.605
MAPE		**8.683**			**8.866**			**8.447**		

The low values of MAPE indicate the high accuracy of all models. However, it is evident from Table 6 that the ANN model is slightly more accurate (MAPE= 8.441) than both linear (MAPE= 8.683) and non-linear (MAPE= 8.866) regression models. Additionally, the ANN model does not suffer from any stability problem, which undermines the reliability of the non-linear model. Thus, it is concluded that the ANN technique is more suitable for modelling the mark-up process based on the bid mark-up factors used in this study.

Discussion

Correlation analysis conducted on the real bidding situations revealed that some important factors do not have a high correlation with the mark-up size. These factors include:

1. "Experience on similar projects". This is an important factor. However, its effect is accounted for in other factors such as "Confidence in the cost estimate". Contractors can

not be so confident in estimating the cost of a new project without considerable experience in similar ones.

2. "Project size". It is expected that contractors would accept less profit for large projects. On the other hand, large projects imply higher risks and subsequently demand higher mark-up to cover them. This might justify the low correlation between the mark-up and the subjective assessment of this factor.
3. "Availability of equipment required". Generally, construction equipment is readily available in Syria compared to the volume of projects being constructed. Therefore, contractors might not need to worry about this factor ($r = 0.005$). However, it is important to distinguish between the availability of equipment in the market and the availability of owned equipment, which has considerable effect on the mark-up ($r = -0.636$).

These factors have more effect on the "bid/no bid" decision ([3], [4]) than on the mark-up decision. Nevertheless, the developed models proved to be highly accurate in simulating the actual make-up percentages in the validation sample. This high accuracy might be attributed to the small size of the validation sample. The models might not be that accurate for larger samples. The development of ANN models, likewise the non-linear models, involves series of trial and error experiments. But, it is much easier and faster than finding the best non-linear regression equation. The ANN model can be adapted to new changes in the bidding practice by retraining on new bidding situations. On the other hand, the regression model needs to be redeveloped from scratch. Also, the regression models lack the ability to generalise solutions to highly correlated, incomplete, or previously unknown data. The "black-box" feature is the main drawback of ANN models. Additionally, the user needs some knowledge of NeuralWorks development software to be able to take advantage of this model. However, the connection weights of the developed model can be extracted manually and used to develop a user-friendly computer programme.

Conclusion

Through a formal questionnaire survey, the most influential factors characterising the mark-up selection in Syria were identified. Data on one hundred and eleven bidding situations was collected from general Syrian contractors. The size of participating contractors was not taken into account in this study. It is recommended that future research is conduced to study the relationship between mark-up estimation practice and the size of contructor. The collected data is composed of subjective assessments of real life bidding situations in terms of the identified mark-up factors along with the actual mark-up percentage selected in each situation. This data was used to develop and validate linear regression, non-linear regression and ANN mark-up models. The present work has identified the key factors influencing mark-up size in Syrian construction projects, demonstrated the relationship between these factors and the mark-up selection. It has also proved that Artificial Neural Networks techniques may be a more accurate predictive tool in modelling the mark-up decision compared to the regression analysis. Unlike usual practice, this is not only based on the accuracy of predictions but also on the robustness of the compared models.

REFERENCES

Ahmad I. and Minkarah, I. (1987). Optimum mark-up for bidding: a preference-uncertainty trade-off approach. Civil Engineering Systems, 4(4), pp170-174.

Ahmad, I., and Minkarah, I. A. (1988). Questionnaire survey on bidding on construction. Journal of Management in Engineering, ASCE, 4 (3), pp 229-243.

Ahmad, I. (1988). Bidding strategy: Multicriteria decision making approach, thesis presented to the University of Cincinnati, Ohio, in partial fulfilment of the requirement of the degree of Doctor of Philosophy.

Broemser, G.M. (1968). Competitive bidding in the construction industry. PhD Dissertation. Graduate School of Business, Stanford University, California, USA.

Benjamin, N.B.H. and Meador, R.C. (1979). Comparison of Friedman and Gates Competitive Bidding Models, Journal of Construction Division, 105 (C01), pp 25-40.

Boussabaine, A. H. (1995). The use of artificial neural networks in construction management: a review. Construction Management and Economics, (14), pp 427-436.

Couzens, A., Skitmore, M., Thorpe, T, and McCaffer, R. (1996). Strategic information requirements in contract bidding: Framework for development of a decision support system. Civil Engineering Systems, Vol. 13, pp. 121-139.

Drew, D. and Skitmore, M. (1997). The effect of contract type and size on competitiveness in bidding. Construction Management and Economics, Vol. 15, pp 469-489.

Dikmen, I., Birgoul, M.T. and Gur, A.K. (2007). A case-based decision support tool for bid mark-up estimation of international construction projects. Automation in Construction. Vol.17, pp 30-44.

Egemen, M. and Mohamed, A.N. (2008). SCBMD: A knowledge-based system software for strategically correct bid/no bid and mark-up size decisions. Automation in Construction. Vol.17, pp 864-872.

ElSawy, I., Hosny, H. and Razek, M.A. (2011). A neural network model for construction site overhead cost estimation in Egypt. International Journal of Computer Science Issues, 8 (3), pp 273-283.

Friedman, L. (1956). A Competitive Bidding Strategy. Operational Research, 4, pp104-112.

Gates, M. (1976). Gates' bidding model, a Monte Carlo experiment. J. Division, ASCE, 102 (4), pp 669-679.

Gates, M. (1983). A bidding strategy based on ESPE. Cost Engineering, 25 (6), pp 27-35.

Gaareslev, A. (1991). Neural network. Technique tested on classical bidding strategy data set. E&FN Spon, London.

Ioannou, P. G. (1988). Bidding Models: symmetry and states of information. Journal of Management in Engineering, ASCE, 114 (2), pp 214-232.ASCE, 102 (4), pp 669-679.

Laryea, S. and Hughes, W. (2011). Risk and Price in the Bidding Process of Contractors. Journal of Construction Engineering and Management, 137 (4), pp.248-258.

Moselhi, O., Hegazy, T., and Fazio, P. (1991). Neural networks as tools in Construction. ASCE Journal of Engineering and Management. 117, 4, pp. 606-625.

Medhi, J. (1992). Statistical Methods: An Introductory Text. New Age International. pp. 53–58.

Polat, G., Baytekin, S. and Eray, E. (2015). Mark-up size estimation in railway projects using the integration of AHP and regression Analysis Techniques. Procedia Engineering, 123, pp 423-431.

Shash, A. (1993). Factors considered in tendering decisions by top UK contractors. Construction Management and Economics, Vol.11, No.2, pp.111-118.

Wanous, M., Boussabaine, A.H. and Lewis, J. (1998). Tendering factors considered by Syrian contractors. ARCOM, 14th annual conference proceedings, Vol. 2, pp 535-544.

Wanous, M., Boussabaine, A.H. and Lewis, J. (2000). Bid/no bid: a parametric solution. Construction Management and Economics. Vol. 18, pp 457-466.

Appendix I: Questionnaire Survey

Factors affecting mark-up size decision in Syria

It is commonly accepted that *mark-up size* decisions are highly complex because they are liable to be affected by many internal and external factors.

The main objective of this study is to identify these factors aiming to develop a decision support tool for making mark-up size decisions.

(Part one)
General Information

The typical Type(s) of the projects you deal with (Please add a √ as appropriate):

- ◯-Building: → ◯-Housing, ◯-Industrial, ◯-Educational, ◯-Offices
- ◯-Pipelines: → ◯-Drinking Water pipes, ◯-Waste pipes.
- ◯-Dams.
- ◯-Roads. ◯-Others (please specify): ....

1. The typical size of the projects you deal with (In Sy.P). (Please tick as appropriate):

◯--Less than (10,000,000),
◯--(10,000,000 – 30,000,000),
◯-(30,000,000 – 50,000,000),
◯--(50,000,000 – 70,000,000),
◯--(70,000,000 – 100,000,000),
◯---More than 100,000,000.

2. Minimum capital required to bid for a new project: Percentage of project size, ... %

3. Current degree of competition. (Please tick as appropriate):

Very low,	Low,	Medium,	High,	Very high.
◯	◯	◯	◯	◯

4. Average number of competitors: (Please tick as appropriate)

3 or Less	4-7	8-10	11 or More
◯	◯	◯	◯

5. Method(s) used in making bid/no bid and mark-up decisions:

◯-Statistical/ Mathematical.
◯-Subjective Judgement.
◯- Other methods (please specify):...............................
....

7-Number of projects tendered for per year: .... Project(s).

-Number of Project(s) obtained per year: ....project(s).

8- Years of experience: ....*year(s).*

(Part Two)

- Please rate the importance of each factor in making mark-up size decisions by circling the appropriate score out of 6.

> ***Where:***
> **0**: No importance **1**: Very Low **2**: Low **3**: Medium **4**: High **5**: Very high **6**: Extreme Importance

Factor Name	None 0	V.L. 1	L. 2	M. 3	H. 4	V.H. 5	E.I 6
	Important Score						
*- Example ~~~~~~~~~~~~~~~~~~~~~~.	0	1	2	3	4	5	6
1- Project Size.	0	1	2	3	4	5	6
2- Sufficiency of the original project duration estimated by the client.	0	1	2	3	4	5	6
3- Sufficiency of the original price estimated by the client.	0	1	2	3	4	5	6
4- Project location.	0	1	2	3	4	5	6
5- Availability of time for tendering.	0	1	2	3	4	5	6
6- Expected project cash flow.	0	1	2	3	4	5	6
7- Degree of buildability.	0	1	2	3	4	5	6
8- Confidence of the cost estimate.	0	1	2	3	4	5	6
9- Rigidity of specifications.	0	1	2	3	4	5	6

0: No importance	**1**: Very low	**2**: Low	**3**: Medium	**4**: High	**5**: Very high	**6**: Extreme importance

Factor Name	Important Score					
10- Risks expected.	0	1	2	3	4	5	6
11- Local climate.	0	1	2	3	4	5	6
12- Local customs.	0	1	2	3	4	5	6
13- Public exposure.	0	1	2	3	4	5	6
14- Site accessibility.	0	1	2	3	4	5	6
15- Site clearance of obstructions.	0	1	2	3	4	5	6
16- Degree of hazard.	0	1	2	3	4	5	6
17- Proportions to be subcontracted.	0	1	2	3	4	5	6
18- Current work load.	0	1	2	3	4	5	6
19- Experience in similar projects.	0	1	2	3	4	5	6
20- Relation with and reputation of the owner.	0	1	2	3	4	5	6
21- Past profit in similar projects.	0	1	2	3	4	5	6
22- Availability of qualified staff.	0	1	2	3	4	5	6
23- Availability of labour/ additional supervisory persons required.	0	1	2	3	4	5	6
24- Availability of capital required.	0	1	2	3	4	5	6
25- Relations with other contractors and suppliers.	0	1	2	3	4	5	6
26- The project's geological study.	0	1	2	3	4	5	6

0: No importance **1**: Very low **2**: Low **3**: Medium **4**: High **5**: Very high **6**: Extreme importance							
Factor Name	Important Score						
27- Fulfilling the to-tender conditions imposed by the client.	0	1	2	3	4	5	6
28- Availability of materials required.	0	1	2	3	4	5	6
29- Availability of equipment owned by the contractor	0	1	2	3	4	5	6
30- Availability of additional equipment required.	0	1	2	3	4	5	6
31- Fluctuation in labour/ material prices.	0	1	2	3	4	5	6
32- Availability of other projects.	0	1	2	3	4	5	6
33- Expected degree of competition)	0	1	2	3	4	5	6
34- Competence of the expected competitors	0	1	2	3	4	5	6

- **Please add any missing factors.**

35-	0	1	2	3	4	5	6
36-	0	1	2	3	4	5	6
37-	0	1	2	3	4	5	6
38-	0	1	2	3	4	5	6